Space-filling curves and related functions

by

Stephen M. Buckley

In this paper, we shall investigate several questions related to space-filling curves. We start with a question whose answer has been known (although not widely known, it would appear) for rather a long time.

Question 1. *Do there exist continuous functions* $f : [0, 1] \to \mathbf{R}$ *which take on each of uncountably many values uncountably often?*

The answer is "yes"; in fact the first component of any space-filling curve (Peano curve) is such a function. A recent rather simple example of such a curve can be found in [8]; for more information on space-filling curves, the reader should consult [7].

Here we shall give two rather different methods of constructing examples of functions answering our question. Some examples using the first construction have zero derivative almost everywhere, while the second construction always leads to nowhere-differentiable examples. We use the notation $f^{-1}(y)$ to denote the set of all pre-images of y.

We begin with the "digit coding" construction. The example we give maps the unit interval onto itself and takes on all its values uncountably frequently. First note that any function continuous on a closed subset S of $[0, 1]$ (with respect to the subspace topology) can be extended to a function continuous on the whole interval simply by filling in the omitted open intervals with continuous interpolating functions (for instance we can "join up the dots" in a linear fashion, and extend the function in a constant fashion at an omitted end-segment). Thus it suffices to find a continuous function $f : S \to [0, 1]$ such that $f^{-1}(y)$ is uncountable for all $y \in [0, 1]$

Suppose $k \geq 3$ is an integer. Any x in $[0, 1]$ can be written in at least one base-k expansion $x = 0.x_1x_2x_3x_4x_5x_6\ldots$. Let S be the (closed) set of all $x \in [0, 1]$ whose even-subscripted digits are all less than $k - 1$. If $x \in S$, then $f(x)$ is defined as the number whose base-k expansion is obtained by removing the even-subscripted digits of x, i.e. $f(x) = 0.x_1x_3x_5\ldots$. Clearly $f^{-1}(y)$ is uncountable (it has the cardinality of the continuum) for every $y \in [0, 1]$. We are left with showing that f is continuous on S. This is also easy but, as we shall use similar arguments several times later, let us give a little more detail this first time. Suppose $f(x) \neq f(y)$ and the first difference occurs at the m'th digit in their base-k expansion, and so $|f(x) - f(y)| \leq k^{-m+1}$. It follows that the $(2m-1)$'st base-k digit of x and y must differ, and the fact that $x, y \in S$ now implies that $|x - y| \geq k^{-2m}$. Thus we are done.

The function f constructed above does not have the bonus property of having zero derivative almost everywhere, but a small adjustment fixes this: one simply uses singular continuous functions to interpolate on the omitted segments rather than linear ones. For instance if (a, b) is one of the omitted intervals, define

$$f(a + t(b - a)) = f(a) + (f(b) - f(a))h(t), \qquad \text{for all } 0 < t < 1,$$

where $h : [0,1] \to [0,1]$ is any increasing continuous function such that $h(0) = 0$, $h(1) = 1$, $h'(x) = 0$ for almost all $x \in (0,1)$ (a basic example of such a function, due to Lebesgue, is described in [4, p.113]).

A slight variant of the above construction leads to a Peano curve. For example in two dimensions, we first define S' to be the (closed) set of all $x \in [0,1]$ *all* of whose digits are less than $k-1$. If $x \in S'$, then $f_1(x)$ and $f_2(x)$ are defined as the numbers whose base-$(k-1)$ expansions are given by $f_1(x) = 0.x_1x_3x_5\ldots$ and $f_2(x) = 0.x_2x_4x_6\ldots$. We extend f_1 and f_2 to all of $[0,1]$ as before. It follows readily that $F = (f_1, f_2)$ is continuous from $[0,1]$ onto $[0,1] \times [0,1]$.

Our second construction uses lacunary functions. Examples of this type are easy to write down, but proving that they have the required properties requires some effort. A typical example is $f(x) = \sum_{j=0}^{\infty} 4^{-j} \cos(100^j x)$. More generally we have the following result:

Theorem 1. *Suppose $f(x) = \sum_{j=0}^{\infty} a_j \cos(b_j x)$, where $a_j, b_j > 0$, and $a_{j+1} < a_j/4$, for all $j \geq 0$. Suppose further that there exists an integer j_0 such that if $j \geq j_0$ then $a_{j+1}b_{j+1} > 6\pi a_j b_j$. Then f is continuous and takes on all values in the interior of its range uncountably frequently.*

Before proving Theorem 1, let us make a few remarks. First of all, f is clearly uniformly continuous on $\mathbf{R}$ since $a_{j+1} < a_j/4$. For the same reason, f takes on all values between $-2a_0/3$ and $2a_0/3$, so Theorem 1 asserts that f takes on uncountably many values uncountably frequently. The numbers b_{j+1}/b_j do not have to be integers, so f may not be periodic.

Proof of Theorem 1: Let us write $s_n(x) = \sum_{j=0}^{n} a_j \cos b_j x$, $r_j = 4a_{j+1}/3$, and $L_j = \pi/b_j$. We fix an arbitrary point c in the interior of the range of f and let n_0 be chosen so large that $n_0 \geq j_0$ and that f takes on all values in the interval $[c - 2r_{n_0}, c + 2r_{n_0}]$. We also assume n_0 is chosen so large that $\sum_{j=0}^{n} a_j b_j \leq 2a_n b_n$ for all $n \geq n_0$; this is possible by the geometric growth assumption on $a_j b_j$.

Given $n \geq 0$, we call an open interval I a *level-n trap* if the values of $s_n - c$ at the endpoints of I are of opposite sign and larger than r_n in absolute value. Note that a level-n trap contains roots of $f - c$ and $s_m - c$ for all $m \geq n$.

There exists a level-n_0 trap, call it (u, v), since we can solve the equations $f(u) = c - 2r_{n_0}$ and $f(v) = c + 2r_{n_0}$. Writing $I_{n_0,1} = (u, v)$, we shall construct a nested binary tree of level-n traps $I_{n,k}$, $n \geq n_0$, $1 \leq k \leq 2^{n-n_0}$, all contained in $I_{n_0,1}$. In fact, each $I_{n,k}$ will contain the closure of two disjoint traps $I_{n+1,j}$ and the length of $I_{n,k}$ will tend to 0 as $n \to \infty$. By continuity, any nested sequence of intervals $(I_{n,k_n})_{n=n_0}^{\infty}$ extracted from this tree specifies a unique root of $f - c$ (the single point which is in the intersection of the I_{n,k_n}'s). This root cannot be at the endpoint of any of the containing traps (since they are compactly nested) and so different sequences of intervals lead to different roots. Thus $f^{-1}(c)$ is uncountable as required.

Assume inductively that we have already defined $I = I_{n,k}$. We must prove the existence of two disjoint level-$(n+1)$ traps contained within it. Suppose $x \in I$ is a root of $s_n - c$. $|s_n'|$ is bounded by $2a_n b_n$ and so, if $d_n = a_{n+1}/(4a_n b_n)$, then s_n differs from c by less than $a_{n+1}/2$ on the interval $[x - d_n, x + d_n]$. Since I is a level-n trap, it must contain $[x - d_n, x + d_n]$.

By hypothesis, $3L_{n+1} < 2d_n$. Thus, $[x - d_n, x + d_n]$ contains subintervals of the form $[mL_n, (m+1)L_n]$ for two consecutive values of m. Since $a_{n+1}\cos(b_{n+1}x)$ takes on the values

$\pm a_{n+1}$ at the endpoints of such subintervals, it follows that the interiors of these subintervals are the required level-$(n+1)$ traps, and so we are done. ∎

The constants $1/4$ and 6π are only convenient values for the proof and are far from sharp. If one examines the proof one sees that the choice of the former constant affects the latter but, even if we leave $1/4$ unchanged, 6π can be replaced by, say, $18\pi/7$ if we estimate things a little more carefully. In fact, since $18\pi/7$ is a little larger than 8, we can choose n_0 so large that $|s'_n|$ is less than sa_nb_n, where $s = 1/(1 - 7/18\pi) < 8/7$. Thus if $d_n = ta_{n+1}/(sa_nb_n)$ for any $t < 2/3$, then s_n differs from c by less than ta_{n+1} on the interval $[x - d_n, x + d_n]$. Choosing t close enough to $2/3$, we have $t/s > 7/12$, and so $3L_{n+1} < 2d_n$ as before.

Incidentally, it follows from the proof of Theorem 1 that any f considered here exhibits the Weierstrass property of being nowhere differentiable. In fact, the variation of f on $I_j = [x - L_j, x + L_j]$ is at least $2a_j$, ensuring that for some $y \in I_j$,

$$|f(y) - f(x)|/|y - x| \geq a_jb_j/\pi \to \infty (j \to \infty).$$

This non-differentiability result is much less sharp, however, than that of Hardy [3], who proved that $\sum a^n \cos b^n \pi x$ is a continuous nowhere-differentiable function whenever $0 < a < 1$, $ab \geq 1$. This suggests the following question:

Question 2. *Do all continuous nowhere-differentiable lacunary series take on uncountably many values uncountably frequently?*

I have no answer to this question, but I would be rather surprised if it were true; perhaps more likely to be true is the conjecture that $\sum a^n \cos b^n \pi x$ takes on a whole interval of values uncountably frequently whenever $0 < a < 1$, $ab > 1$ (since here we have some "room to manoeuvre").

Before going on to our next question, let us introduce some terminology that we need here and later. Given $0 < t < 1$, we say $f : [0, 1] \to \mathbf{R}^n$ is *t-Hölder continuous* if, for some $C > 0$,

$$|f(x) - f(y)| \leq C|x - y|^t, \qquad \forall\, 0 \leq x, y \leq 1. \tag{1}$$

In the case $t = 1$, we instead say that f is *Lipschitz (continuous)*, or C-Lipschitz if we wish to specify the constant.

Question 3. *Do there exist Lipschitz functions $f : [0, 1] \to \mathbf{R}$ which take on each of uncountably many values uncountably often?*

The answer to Question 3 is again "yes," although examples like the previous ones fail because $f^{-1}(x)$ must be finite almost everywhere (see Theorem 2 below). Instead we first define f on $S \subset [0, 1]$, the closed set of numbers whose decimal expansion can be written using only the digits 0, 2, 7, and 9. For these numbers, the decimal expansion of $f(x)$ is calculated from that of x by changing all 2's to 0's, and all 7's to 9's (and so $f^{-1}(x)$ is uncountable for every x whose decimal expansion involves only 0's and 9's). We define f at all other values by linear interpolation. We leave to the reader the task of verifying that the resulting function f satisfies the Lipschitz condition $|f(y) - f(x)|/|y - x| \leq 2$ on S (and hence on $[0, 1]$).

For any exponent $t < 1$, one can construct a t-Hölder continuous $f : [0,1] \to \mathbf{R}$ which takes on all values in an interval uncountably frequently. Our first digit-coding example f is an example for $t = 1/2$. This construction is easily modified to handle any $t < 1$. First let S to be the set of $x \in [0,1]$ for which the base-k expansion has no digit equal to $k-1$ in any position whose subscript is divisible by a fixed integer $m > 1$. We define f on S by deleting all digits whose subscript is divisible by m, and extend f using linear interpolation. Then f is t-Hölder continuous for $t = (m-1)/m$, and $f^{-1}(x)$ is uncountable for all $x \in [0,1]$.

The following theorem shows how different things are for Lipschitz functions. This result is a special case[1] of a more general result concerning Lipschitz maps between metric spaces (see [2, Corollary 2.10.11]), but we give a short proof here for completeness.

Proposition 2. *If $f : [0,1] \to \mathbf{R}$ is C-Lipschitz and $N : \mathbf{R} \to [0,\infty]$ is the cardinality of $f^{-1}(x)$, then $\int_{\mathbf{R}} N(x)\,dx \le C$. Consequently, $N(x)$ is finite almost everywhere.*

Proof: For all $j > 0$, let Δ_j be the collection of dyadic intervals of the form $[(k-1)2^{-j}, k2^{-j})$, for $1 \le k < 2^j$, and $[1-2^{-j}, 1]$. Note that $(\Delta_j)_{j=1}^{\infty}$ is a nested sequence of partitions of $[0,1]$. Let $N_j(x)$ be the number of intervals $f(I)$, $I \in \Delta_j$, which contain x. Using the properties of Δ_j, we see that for each $x \in \mathbf{R}$, $N_j(x)$ is an increasing function of j which tends to $N(x)$ as $j \to \infty$. Furthermore, it is clear that

$$\int_0^1 N_j(x)\,dx = \sum_{I \in \Delta_j} |f(I)| \le \sum_{I \in \Delta_j} C|I| = C,$$

where $|I|$ and $|f(I)|$ denote the lengths of the intervals I and $f(I)$. An appeal to Lebesgue's Monotone Convergence Theorem finishes the proof. □

Question 4. *Does there exist a function f from $[0,1]$ onto $U \equiv [0,1] \times [0,1]$ which is t-Hölder continuous for some $t \ge 1/2$?*

Question 4, like Question 3, is inspired by a shortcoming in the earlier examples: our base-k Peano curve F is t-Hölder continuous for $t = (\log(k-1))/(2\log k)$, thus providing examples for all $t < 1/2$. The following theorem answers Question 4.

Theorem 3. *There exist Peano curves $F : [0,1] \to U$ which are t-Hölder continuous for $t = 1/2$, but no such curve is t-Hölder continuous for $t > 1/2$.*

Proof: We first examine the case $t > 1/2$. The *Minkowski dimension* of a compact subset E of $\mathbf{R}^n$ is defined by

$$\mathcal{M}\text{-dim } E = \sup\{s \ge 0 : \limsup_{r \to 0} H_s(E,r) = \infty\},$$

where $H_s(E,r)$ is the *α-dimensional Minkowski precontent*, defined as kr^s, where k is the minimum number of balls of radius r required to cover E. These concepts, and the related concept of Hausdorff dimension, are discussed at greater length in [5] and [1]. We shall need only the easily proven fact that any compact $E \subset \mathbf{R}^n$ of positive measure has Minkowski dimension n.

[1] I would like to thank P. Hajłasz for pointing this out to me.

Also noteworthy, although not needed by us, is the obvious fact that the Minkowski dimension of a set is greater than or equal to its Hausdorff dimension.

Suppose that $F : [0,1] \to U$ satisfies (1) for some $t > 1/2$. We claim that the Minkowski dimension of $F([0,1])$ is at most $1/t$ (and hence the range of F cannot be all of U). To see this note that the image of any interval $[i/k, (i+1)/k]$ is contained in a ball of radius C/k^t about $f(i/k)$. Thus $H_{1/t}(F([0,1]), C/k^t) \leq C^{1/t}$, and our claim follows easily.

We next construct the required 1/2-Hölder continuous Peano curve. The base-3 example I shall give is the same as Peano's original example of a space-filling curve[2] [6]. The basic idea is simple: we can "almost" get the solution by "chopping" x into its base-k digits, allocating them one at a time to be the next base-k digit of either $f_1(x)$ or $f_2(x)$. This certainly gives a space-filling function but it is not 1/2-Hölder continuous (or even continuous) because of the following phenomenon: if $y = 0.y_1y_2\dots y_n\dots$ in base-k, where $y_m \neq 0$ and $y_n = 0$ for all $n > m$, and if m is odd (even) then the left- and right-hand limits for f_2 (respectively f_1) at y are different. The way out of this problem is fairly clear: we allocate digits one at a time to $f_1(x)$ and $f_2(x)$ but introduce a parity effect to compensate for these discontinuities. We describe this process for base 3 where it is most easily done.

To avoid problems caused by non-unique expansions, we define functions $A : [0,1] \to S$ and $B : S \to [0,1]$, where S is the set of infinite sequences whose terms are restricted to the set $\{0,1,2\}$. A maps numbers to (one of) their base-3 expansions, and B maps $(x_1, x_2, x_3, \dots)$ to the number with base-3 expansion $0.x_1x_2x_3\cdots$. We shall write values of these functions in the form Ay and Bx. Whenever $x \in S$, we denote its i'th term by x_i.

Let $G = (g_1, g_2) : S \to S \times S$ be defined by $G(x) = (u, v)$ where

$$u_k = \begin{cases} x_{2k-1}, & \text{if } \sum_{i=1}^{k-1} x_{2i} \text{ is even} \\ 2 - x_{2k-1}, & \text{if } \sum_{i=1}^{k-1} x_{2i} \text{ is odd,} \end{cases}$$

$$v_k = \begin{cases} x_{2k}, & \text{if } \sum_{i=1}^{k} x_{2i-1} \text{ is even} \\ 2 - x_{2k}, & \text{if } \sum_{i=1}^{k} x_{2i-1} \text{ is odd.} \end{cases}$$

We now define $F(y) = (Bg_1(Ay), Bg_2(Ay))$ whenever $y \in [0,1]$.

Clearly F has range U. We are left with showing that F is 1/2-Hölder continuous. A simple case-by-case argument reveals that F is independent of the choice of A (for example, $G(0,2,2,2,\dots) = G(1,0,0,0,\dots)$). Whenever $x \in S$, $G(x) = (u,v)$, let us call $u_1, v_1, u_2, v_2, u_3, v_3, \dots$ the *standard order* of the terms of $G(x)$.

Suppose $x, y \in S$ and $Bx < By$. Let us assume that the first term of $G(x)$ which differs from the corresponding digit of $G(y)$, using the standard order, is the m'th term of the $g_2(y)$ (if the first difference is in $g_1(y)$, a similar argument applies). Then $|F(Bx) - F(By)| \leq 3^{-m+1}$. If $|Bx - By| \geq 3^{-2m}$, we are done, so we may assume $|Bx - By| < 3^{-2m}$. But then, if there is some $0 < j < 2m$ such that $x_i = y_i$ if $i < j$ and $x_j \neq y_j$, we must have $x_j + 1 = y_j$ and, whenever $j < i \leq 2m$, $x_i = 2$ and $y_i = 0$. This forces the m'th digit of the second coordinates of $F(Bx)$ and $F(By)$ to be equal, contrary to assumption. The only remaining possibility is that $x_i = y_i$

[2] I would like to thank the editor for sending me a copy of this paper.

if $i < 2m$ and $y_{2m} = x_{2m} + 1$. In this case, $Bx \le Bz \le By$, where $z_i = y_i$ for $i \le 2m$ and $z_i = 0$ if $i > 2m$. If $z \ne y$ and the j'th term is the first term where they differ, then it is clear that

$$By - Bx \ge By - Bz \ge 3^{-j}, \qquad |F(y) - F(z)| \le 3^{1-j/2}. \tag{2}$$

Next let z' be the sequence defined by $z_i = x_i$ for $i \le 2m$ and $z_i = 2$ if $i > 2m$, so that $Bz = Bz'$. If $z' \ne x$ and the k'th term is the first term where they differ, then it is again clear that

$$By - Bx \ge Bz - Bx \ge 3^{-k}, \qquad |F(z) - F(x)| \le 3^{1-k/2}. \tag{3}$$

Putting (2) and (3) together, we get the desired Hölder continuity. ∎

Our previous argument actually implies that there are no Peano curves $f : [0,1] \to [0,1]^n$, $n \ge 2$, which are t-Hölder continuous for $t > 1/n$. The construction for $t = 1/2$ also generalises to give an n-dimensional Peano curve which is $1/n$-Hölder continuous in the higher dimensional setting: again using a base-3 expansion, we "deal out" the digits one at a time to each of the n coordinates, replacing each "0" by "2" and vice versa whenever the sum of the digits previously dealt to the other coordinates is odd. We leave the verification of $1/n$-Hölder continuity to the reader.

Question 5. *Does there exist a map G from the unit square $U = [0,1] \times [0,1]$ to U such that the image of any non-trivial line segment in U has non-empty interior?*

We give a couple of methods for constructing such a map G. The map $A(x,y) = F(x)$, where $F = (f,g)$ is the Peano curve defined earlier, has this property on all non-vertical lines. Defining $B : U \to U$ by $B(x,y) = ((x+y^2)/2, y)$, $G = A \circ B$ has the desired property (since if $L \subset U$ is a non-trivial line segment, the x-projection of $B(L)$ is also a non-trivial line segment).

One might feel that the previous method is not completely satisfactory since we have simply "hidden" the straight lines. Our second method, has the advantage that it produces a function G for which the image of $G \circ \gamma$ has non-empty interior whenever γ is a non-trivial C^1 path in U. First let $F_k = (f_k, g_k)$ to be our old base-k Peano curve F. F_k is t-Hölder continuous for $t = (\log(k-1))/(2\log k)$ but not for any larger t; in fact, it is easily seen that for any n, the image of any interval of length $1/k^{2n}$ under F_k is contained in a square of length $(k-1)^{-n+1}$ and contains a square of length $(k-1)^{-n-1}$.

We claim that $G(x,y) = F_i(x) + F_j(y)$ is a function of the type we require for any $3 \le i < j$. We shall content ourselves here with sketching the proof. Clearly images of vertical line segments have non-empty interior. If γ is not a vertical line, then we need only look in the vicinity of a single point (x_0, y_0) on γ where the tangent line is non-vertical. In this case, one expects everything to work out since on any sufficiently small neighbourhood of γ (dependent on the slope of the tangent line), the variation in F_i is much larger than the variation in F_j. To make this idea rigorous, assume $G(x_0, y_0) = (u_0, v_0)$. We solve the equation $G(x,y) = (u,v)$ for all (u,v) sufficiently near (u_0, v_0) by an iterative method. Having found the approximate solution (x_k, y_k), we find $(x_{k+1}, y_{k+1}) \in I$, a nearby point on the curve for which $F_i(x_{k+1}) + F_j(y_k) = (u,v)$. With this hint, we leave the details to the reader.

Acknowledgements. A couple of variations of Question 1 were posed to me by Finbarr Holland. Piotr Hajłasz asked me Questions 3 and 5 as well as suggesting several useful references. I would like to thank them both for useful related conversations.

References

[1] K.J. Falconer, The geometry of fractal sets, Cambridge University Press, Cambridge, 1985.

[2] H. Federer, Geometric measure theory, Springer-Verlag, Berlin, 1969.

[3] G.H. Hardy, "Weierstrass's non-differentiable function," Trans. Amer. Math. Soc. 17 (1916), 301–325.

[4] E. Hewitt and K. Stromberg, Real and abstract analysis, Springer-Verlag, Berlin, 1965.

[5] O. Martio and M. Vuorinen, "Whitney cubes, p-capacity, and Minkowski content," Expo. Math. 5 (1987), 17–40.

[6] G. Peano, "Sur une courbe, qui remplit tout une aire plaine," Math. Ann. 36 (1890), 157.

[7] H. Sagan, Space-filling curves, Springer-Verlag, Berlin, 1994.

[8] L. Wen, "A space filling curve," Amer. Math. Monthly 90 (1983), 283.

Stephen M. Buckley, Department of Mathematics, St. Patrick's College, Maynooth, Co. Kildare, Ireland.
e-mail: sbuckley@maths.may.ie